

МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ НАЦІОНАЛЬНИЙ ТЕХНІЧНИЙ УНІВЕРСИТЕТ УКРАЇНИ «КИЇВСЬКИЙ ПОЛІТЕХНІЧНИЙ ІНСТИТУТ ІМЕНІ ІГОРЯ СІКОРСЬКОГО» ФІЗИКО-ТЕХНІЧНИЙ ІНСТИТУТ

ЛАБОРАТОРНА РОБОТА №1

3 дисципліни «Криптографія»

Виконали:

студенти 3 курсу ФТІ групи ФБ-93 Абдуллаєва Есміра Шовак Мирослав

Викладач:

Селюх П. В.

Мета роботи: засвоєння понять ентропії на символ джерела та його надлишковості, вивчення та порівняння різних моделей джерела відкритого тексту для наближеного визначення ентропії, набуття практичних навичок щодо оцінки ентропії на символ джерела.

Завдання

- 1. Написати програми для підрахунку частот букв і частот біграм в тексті, а також підрахунку Н1 та Н2 за безпосереднім означенням. Підрахувати частоти букв та біграм, а також значення Н1 та Н2 на довільно обраному тексті російською мовою достатньої довжини (щонайменше 1Мб), де імовірності замінити відповідними частотами. Також одержати значення Н1 та Н2 на тому ж тексті, в якому вилучено всі пробіли.
 - 2. За допомогою програми CoolPinkProgram оцінити значення H (10), H (20), H (30).
- 3. Використовуючи отримані значення ентропії, оцінити надлишковість російської мови в різних моделях джерела.

Хід роботи та опис труднощів:

Під час роботи, з огляду на вимоги форматування тексту, та різних операцій над ним, було вирішено використовувати мову python. У якості вхідного тексту була обрана книга "Зелена миля". Форматування вхідного тексту не викликало значних труднодщів . За допомогою стандартних функцій ми змогли позбутися пробілів, твердого знаку, змінити всі великі букви на малі. За допомогою регулярного виразу змогли досягнути зчитування з файлу тільки алфавіту. Для зручнішого використання програми , в залежності від завдання, було вирішено зробити лаконічне меню з 3 можливостями: зчитування файлу з пробілами і виконання операцій над знаками, зчитування файлу без пробілів, аналогічно було проведено для біграм. Сортування монограм ми досягли за допомогою питру. Труднощі виникли при створення функції для зчитування біграм. Було прийнято рішення створити масив букв, і вже до значень цього масиву додавати як аргумент фукції ще одну букву по черзі і потім рахувати утворені біграми у списку, який ми отримали за допомогою line = [one[k:k+2]] for k in range(0, len(one), 2)](випадок з неперехресними біграмами) та line = [one[k:k+2]] for k in range(len(one))](випадок з перехресними). Вивід реалізовували у вигляді квадратної матриці, для презентабельного вигляду ми відформатували значення наших Н2 до трьох знаків після коми.

Код програми

```
P = 0
import math
import re
                                                                    P = -math.log2(fr)
import numpy as np
                                                                 H = fr * P
open ("/Users/esmira.23/Desktop/KPI/Зкурс/Крипт
                                                                 save1.append("%.3f" % fr) # or H
a/1.txt", "r")
                                                                 save2.append(H)
data = file.read().replace(""",""")
                                                             print(alp[a], "
                                                                              ".join(map(str, save1)))
one = re.sub(r"[^a-s]+", " ",
data.lower()).replace("b","b").replace("ë","e"
                                                             save1.clear()
                                                         if ans == "1":
                                                             # with spaces
# m=32 - кількість букв алфавіту
                                                             number = len(one)
alphabet =
                                                             func (one, number)
['a','6','B','r','Д','e','ж','3','N','Й','K','
                                                             x = sum(save)
print("H1: ", x)
R = 1 - x / math.log2(32)
                                                             print("R for H1 with spaces: ", R)
alphabet1 =
['a','б','в','г','д','е','ж','з','и','й','к','
                                                             # without spaces
print("\n")
                                                             save.clear()
arr =
                                                             save1.clear()
[0,1,2,3,4,5,6,7,8,9,10,11,12,13,14,15,16,17,1
                                                             save2.clear()
                                                             two = one.replace(" ", "")
8,19,20,21,22,23,24,24,26,27,28,29,30,31]
arr1 =
                                                             number1 = len(two)
[0,1,2,3,4,5,6,7,8,9,10,11,12,13,14,15,16,17,1
                                                             func(two, number1)
8,19,20,21,22,23,24,24,26,27,28,29,30]
                                                             x = sum(save)
                                                             print("H1: ", x)
R = 1 - x / math.log2(32)
save = []
save1 = []
save2 = []
                                                             print("R for H1 without spaces: ", R)
                                                         elif ans == "2":
                                                             # не перетинаються
# with spaces
print("1. Leters with and without spaces\n2.
Bigram without crossing\n3. Bigram with
                                                             number = len(one)/2
print(" ", " ".join(alphabet))
crossing")
ans = input()
                                                             for item in arr:
def func(text, numb):
                                                               bigram(item, alphabet, number)
    for i in alphabet:
                                                             x = sum(save2) / 2
       letter = text.count(i)
                                                             print("H2 with spaces: ", x)
        fr = letter / numb
                                                             R = 1 - x / math.log2(32)
        if letter == 0:
                                                             print("R for H2 with spaces: ", R)
            P = 0
                                                             save2.clear()
                                                             print ("\n")
        else:
           P = -math.log2(fr)
                                                             # without spaces
                                                             one = one.replace(" ", "")
        H = fr * P
                                                             number1 = len(one)/2
print(" ", "
        save.append(H)
                                                                                  ".join(alphabet1))
        savel.append(i)
        save2.append("%.4f" % fr) # or H
                                                             for item in arr1:
                                                             bigram(item, alphabet1, number1)
x = sum(save2) / 2
    a = np.column_stack([save1, save2])
    b = sorted(a, key = lambda x: x[1])
                                                             print("H2 without spaces: ", x)
R = 1 - x / math.log2(32)
    c = "\n".join(map(str,b))
    print(c)
                                                             print("R for H2 without spaces: ", R)
                                                         elif ans == "3":
def bigram(a, alp, numb):
                                                             # перетинаються
                                                             # with spaces
    line = [one[k:k + 2] for k in range(0,
len(one), 2)]
                                                             number = len(one)-1
print(" ", " ".join(alphabet))
    for i in alp:
        r = [alp[a] + i]
                                                             for item in arr:
                                                             bigram2(item, alphabet, number)
x = sum(save2) / 2
        bigram1 = line.count(r[0])
        fr = bigram1 / numb
        if bigram1 == 0:
                                                             print("H2 with spaces: ", x)
            P = 0
                                                             R = 1 - x/math.log2(32)
                                                             print("R for H2 with spaces: ", R)
            P = -math.log2(fr)
                                                             save2.clear()
        H = fr * P
                                                             print("\n")
        save1.append("%.3f" % fr) # or H
                                                              # without spaces
                                                             one = one.replace(" ", "")
        save2.append(H)
                                                             number1 = len(one)-1
print(" ", " ".join(alphabet1))
    print(alp[a], " ".join(map(str, save1)))
    save1.clear()
                                                             for item in arr1:
def bigram2(a, alp, numb):
                                                               bigram2(item, alphabet1, number1)
    line = [one[k:k+2] for k in
                                                             x = sum(save2) / 2
                                                             print("H2 without spaces: ", x)
range(len(one))]
                                                             R = 1 - x / math.log2(32)
    for i in alp:
        r = [alp[a] + i]
bigram1 = line.count(r[0])
                                                             print("R for H2 without spaces: ", R)
                                                         else:
        fr = bigram1 / numb
                                                             print("Wrong input")
        if bigram1 == 0:
```

Значення Н1 для тексту з пробілами: Значення Н1 для тексту без пробілів: ['¢' '0.0022'] ['''0.0000'] ['ц' '0.0024'] ['ф' '0.0026'] ['щ' '0.0026'] ['ц' '0.0029'] ['a' '0.0038'] ['щ' '0.0032'] ['ю' '0.0047'] ['9' '0.0046'] ['x' '0.0073'] ['ю' '0.0057'] ['ш' '0.0074'] ['x' '0.0088'] ['й' '0.0084'] ['ш' '0.0089'] ['ж' '0.0093'] ['й' '0.0101'] ['4' '0.0134'] ['ж' '0.0113'] ['6' '0.0142'] ['4' '0.0161'] ['3' '0.0142'] ['3' '0.0171'] ['ы' '0.0148'] ['6' '0.0172'] ['r' '0.0149'] ['ы' '0.0179'] ['ь' '0.0152'] ['r' '0.0180'] ['я' '0.0191'] ['ь' '0.0184'] ['n' '0.0243'] ['я' '0.0231'] ['n' '0.0294'] ['y' '0.0260'] ['д' '0.0266'] ['y' '0.0314'] ['m' '0.0276'] ['д' '0.0322'] ['k' '0.0303'] ['m' '0.0333'] ['B' '0.0315'] ['k' '0.0366'] ['p' '0.0366'] ['B' '0.0381'] ['c' '0.0425'] ['p' '0.0442'] ['л' '0.0447'] ['c' '0.0514'] ['л' '0.0540'] ['T' '0.0504'] ['u' '0.0527'] ['T' '0.0609'] ['и' '0.0637'] ['H' '0.0536'] ['a' '0.0643'] ['H' '0.0648'] ['e' '0.0674'] ['a' '0.0777'] ['o' '0.0950'] ['e' '0.0815'] [' ' '0.1724'] ['o' '0.1148']

H1: 4.366187862436842

R for H1 with spaces: 0.12676242751263156

H1: 4.474445521025726

R for H1 without spaces: 0.10511089579485478

Значення Н2 з пробілами без перетину:

H2: 3.978541281663438

					_		_		м					_			_							-				_			
a - 0 000	0 001	B	r 0 000	Д	e	× 001	3	И	И	K	л	M	H	0	П 0.001	p	C	T	у	ф	X	ц	0 001	W 0.001	щ	ы	ь ооо	3	0 001	я	0.017
a 0.000						0.001																								0.001	
6 0.001																										0.003				0.801	
в 0.004																															
r 0.001																															
д 0.004																															
e 0.000																															
ж 0.001																															
з 0.006																															
и 0.000																															
й 0.000																															
к 0.008																															
л 0.006																										0.001					
м 0.003																										0.002					
н 0.010																										0.003					
0.000																															
n 0.002																															
p 0.006																															
c 0.001																															
т 0.006																															
y 0.001																															
ф 0.000																															
x 0.001																															
ц 0.001																															
ч 0.002																										0.000					
ш 0.001																															
w 0.001																															
ы 0.000																															
ь 0.000																															
э 0.000						0.000																				0.000				0.000	
ю 0.000																										0.000					
я 0.000																															
					0.004	0.001	0.005	0.009	0.000	0.010	0.002	0.008	0.017	0.011	0.018	0.004	0.016	0.009	0.005	0.000	0.002	0.000	0.006	0.001	0.000	0.000	0.000	0.003	0.000	0.005	0.000
H2 with	spaces:	3.978	35412816	63438																											

R for H2 with spaces: 0.2042917436673124

Найчастіші біграми(текст з пробілами без перетину біграм)												
3 буквою а	_	0,017										
3 буквою б	Ы	0,003										
3 буквою в	0	0,005										
3 буквою г	О	0,007										
3 буквою д	е	0,005										
3 буквою е	_	0,017										
3 буквою ж	е	0,003										
3 буквою з	а	0,006										
3 буквою и	_	0,019										
3 буквою й	_	0,007										
3 буквою к	а	0,008										
3 буквою л	_	0,011										
3 буквою м		0,008										
3 буквою н	а	0,1										
3 буквою о	_	0,022										

3 буквою п	0	0,011
3 буквою р	0	0,006
3 буквою с	Т	0,009
3 буквою т	0	0,014
3 буквою у	_	0,006
3 буквою ф	И	0,001
3 буквою х	_	0,003
3 буквою ц	е	0,001
3 буквою ч	Т	0,004
3 буквою ш	е	0,002
3 буквою щ	е	0,002
3 буквою ы	_	0,005
3 буквою ь	_	0,01
3 буквою э	Т	0,003
3 буквою ю	_	0,002
3 буквою я		0,013
3_	П	0,018

Значення Н2 без пробілів та перетину:

H2: 4.1880373760474665

а	6	В	г	д	е	ж	3	И	й	к	л	М	н	0	п	р	С	т	У	ф	х	ц	ч	ш	щ	ы	ь	3	Ю	я
a 0.000	0.002	0.005	0.001	0.003	0.002	0.001	0.005	0.001	0.001	0.008	0.011	0.004	0.005	0.001	0.003	0.003	0.006	0.006	0.001	0.000	0.001	0.000	0.002	0.001	0.000	0.000	0.000	0.001	0.001	0.002
6 0.001	0.000	0.000	0.000	0.000	0.002	0.000	0.000	0.001	0.000	0.000	0.001	0.000	0.000	0.003	0.000	0.002	0.000	0.000	0.001	0.000	0.000	0.000	0.000	0.000	0.000	0.004	0.000	0.000	0.000	0.001
в 0.005	0.000	0.000	0.000	0.001	0.005	0.000	0.001	0.003	0.000	0.001	0.001	0.000	0.002	0.007	0.001	0.001	0.004	0.001	0.001	0.000	0.000	0.000	0.000	0.001	0.000	0.002	0.000	0.000	0.000	0.000
r 0.001	0.000	0.000	0.000	0.002	0.000	0.000	0.000	0.001	0.000	0.000	0.002	0.000	0.000	0.009	0.000	0.001	0.000	0.000	0.001	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
д 0.005	0.000	0.001	0.000	0.000	0.007	0.002	0.000	0.003	0.000	0.000	0.001	0.000	0.002	0.004	0.000	0.001	0.001	0.000	0.002	0.000	0.000	0.000	0.000	0.000	0.000	0.001	0.001	0.000	0.000	0.000
e 0.000	0.002	0.003	0.005	0.004	0.002	0.001	0.002	0.001	0.002	0.003	0.008	0.005	0.009	0.001	0.003	0.010	0.006	0.007	0.001	0.000	0.001	0.000	0.002	0.001	0.001	0.000	0.000	0.000	0.000	0.001
ж 0.001	0.000	0.000	0.000	0.001	0.004	0.000	0.000	0.002	0.000	0.000	0.000	0.000	0.001	0.001	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
з 0.007	0.000	0.001	0.000	0.001	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.002	0.001	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
и 0.000	0.001	0.004	0.001	0.003	0.002	0.001	0.003	0.002	0.001	0.003	0.007	0.004	0.006	0.002	0.003	0.001	0.005	0.005	0.001	0.000	0.002	0.001	0.002	0.001	0.000	0.000	0.000	0.000	0.000	0.002
й 0.000	0.000	0.001	0.000	0.001	0.000	0.000	0.000	0.001	0.000	0.001	0.000	0.001	0.001	0.000	0.001	0.000	0.001	0.001	0.000	0.000	0.000	0.000	0.001	0.000	0.000	0.000	0.000	0.000	0.000	0.000
к 0.009	0.000	0.000	0.000	0.000	0.001	0.000	0.000	0.004	0.000	0.000	0.001	0.000	0.001	0.010	0.001	0.003	0.001	0.001	0.003	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
л 0.007	0.001	0.001	0.000	0.001	0.006	0.000	0.001	0.008	0.000	0.001	0.001	0.001	0.002	0.008	0.001	0.000	0.003	0.001	0.002	0.000	0.000	0.000	0.001	0.000	0.000	0.001	0.004	0.000	0.001	0.003
м 0.003	0.000	0.001	0.000	0.001	0.005	0.000	0.000	0.004	0.000	0.001	0.000	0.000	0.003	0.006	0.001	0.000	0.001	0.000	0.003	0.000	0.000	0.000	0.000	0.000	0.000	0.002	0.000	0.000	0.000	0.001
н 0.012	0.000	0.001	0.000	0.001	0.011	0.000	0.000	0.007	0.000	0.001	0.000	0.000	0.003	0.012	0.001	0.000	0.001	0.001	0.004	0.000	0.000	0.000	0.000	0.000	0.000	0.003	0.001	0.000	0.000	0.002
0.000	0.005	0.009	0.006	0.006	0.003	0.003	0.002	0.002	0.004	0.004	0.007	0.008	0.012	0.002	0.004	0.007	0.009	0.010	0.001	0.001	0.001	0.000	0.004	0.001	0.000	0.000	0.000	0.001	0.000	0.002
п 0.002	0.000	0.000	0.000	0.000	0.003	0.000	0.000	0.001	0.000	0.000	0.001	0.000	0.000	0.013	0.000	0.007	0.000	0.000	0.001	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
p 0.008	0.000	0.001	0.000	0.001	0.006	0.000	0.000	0.006	0.000	0.001	0.000	0.000	0.001	0.008	0.000	0.001	0.002	0.001	0.005	0.000	0.000	0.000	0.000	0.000	0.000	0.002	0.001	0.000	0.000	0.001
c 0.002	0.000	0.002	0.000	0.001	0.004	0.000	0.000	0.004	0.000	0.004	0.004	0.002	0.002	0.003	0.002	0.000	0.001	0.011	0.001	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.003	0.000	0.000	0.005
т 0.007	0.000	0.003	0.000	0.001	0.005	0.000	0.000	0.004	0.000	0.001	0.000	0.000	0.002	0.018	0.001	0.003	0.002	0.001	0.002	0.000	0.000	0.000	0.000	0.000	0.000	0.002	0.007	0.000	0.000	0.001
y 0.001	0.001	0.001	0.001	0.002	0.000	0.002	0.001	0.001	0.000	0.002	0.002	0.002	0.001	0.001	0.001	0.001	0.002	0.003	0.000	0.000	0.001	0.000	0.001	0.001	0.000	0.000	0.000	0.000	0.001	0.000
ф 0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.001	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.001	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
x 0.001	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.001	0.003	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
ц 0.001	0.000	0.000	0.000	0.000	0.001	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
4 0.002	0.000	0.000	0.000	0.000	0.004	0.000	0.000	0.002	0.000	0.001	0.000	0.000	0.001	0.000	0.000	0.000	0.000	0.005	0.001	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
w 0.001	0.000	0.000	0.000	0.000	0.003	0.000	0.000	0.001	0.000	0.001	0.001	0.000	0.000	0.001	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.001	0.000	0.000	0.000
ш 0.001	0.000	0.000	0.000	0.000	0.003	0.000	0.000	0.001	0.000	0.001	0.001	0.000	0.000	0.001	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.001	0.000	0.000	0.000
ы 0.000	0.000	0.001	0.000	0.000	0.001	0.000	0.000	0.000	0.002	0.000	0.003	0.002	0.001	0.000	0.001	0.000	0.001	0.001	0.000	0.000	0.001	0.000	0.000	0.001	0.000	0.000	0.000	0.000	0.000	0.000
ь 0.000	0.000	0.001	0.000	0.001	0.001	0.000	0.000	0.001	0.000	0.002	0.000	0.001	0.003	0.001	0.001	0.000	0.002	0.001	0.000	0.000	0.000	0.000	0.001	0.001	0.000	0.000	0.000	0.000	0.000	0.001
э 0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.001	0.000	0.000	0.000	0.000	0.000	0.000	0.003	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
ю 0.000	0.000	0.000	0.000	0.001	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.001	0.000	0.001	0.000	0.000	0.000	0.000	0.000
я 0.000	0.001	0.002	0.000	0.002	0.000	0.000	0.001	0.001	0.000	0.001	0.001	0.001	0.002	0.001	0.002	0.000	0.002	0.002	0.001	0.000	0.000	0.000	0.001	0.000	0.000	0.000	0.000	0.000	0.000	0.000
H2 witho	ut spac	es: 4.	1880373	7604746	65																									
D for 110			0 1	/ 0700F0	/700F0/																									

R for H2 without spaces: 0.1623925247905067

Найчастіші біграми(текст без пробілів та перетину біграм)												
3 буквою а	Л	0,011										
3 буквою б	Ы	0,004										
3 буквою в	0	0,007										
3 буквою г	0	0,009										
3 буквою д	е	0,007										
3 буквою е	Н	0,009										
3 буквою ж	е	0,004										
3 буквою з	а	0,007										
3 буквою и	Л	0,007										
3 буквою й	Д	0,001										
3 буквою к	0	0,01										
3 буквою л	0	0,008										
3 буквою м	0	0,006										
3 буквою н	а	0,012										
3 буквою о	Н	0,012										

	_	
3 буквою п	0	0,013
3 буквою р	а	0,008
3 буквою с	Т	0,009
3 буквою т	0	0,018
3 буквою у	Д	0,002
3 буквою ф	ф	0,001
3 буквою х	0	0,003
3 буквою ц	а	0,001
3 буквою ч	Т	0,005
3 буквою ш	е	0,003
3 буквою щ	е	0,003
3 буквою ы	Л	0,003
3 буквою ь	К	0,002
3 буквою э	Т	0,003
3 буквою ю	Ч	0,001
3 буквою я	Д	0,002

Значення Н2 з пробілами та з перетином:

H2: 3.9801036688866813

а	6	В	г	д	е	ж	3	И	й	к	л	м	н	0	п	р	С	т	у	φ	x	ц	ч	ш	щ	ы	ь	э	Ю	я	
a 0.000	0.001	0.002	0.000	0.002	0.001	0.001	0.004	0.000	0.001	0.005	0.009	0.003	0.003	0.000	0.001	0.002	0.003	0.004	0.000	0.000	0.001	0.000	0.001	0.001	0.000	0.000	0.000	0.000	0.001	0.001	0.017
6 0.001	0.000	0.000	0.000	0.000	0.002	0.000	0.000	0.001	0.000	0.000	0.001	0.000	0.000	0.002	0.000	0.002	0.000	0.000	0.001	0.000	0.000	0.000	0.000	0.000	0.000	0.003	0.000	0.000	0.000	0.001	0.000
в 0.004	0.000	0.000	0.000	0.000	0.004	0.000	0.000	0.003	0.000	0.000	0.000	0.000	0.001	0.005	0.000	0.001	0.003	0.000	0.001	0.000	0.000	0.000	0.000	0.000	0.000	0.002	0.000	0.000	0.000	0.000	0.005
r 0.001	0.000	0.000	0.000	0.002	0.000	0.000	0.000	0.001	0.000	0.000	0.002	0.000	0.000	0.007	0.000	0.001	0.000	0.000	0.001	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.001
д 0.004	0.000	0.001	0.000	0.000	0.005	0.002	0.000	0.002	0.000	0.000	0.001	0.000	0.002	0.003	0.000	0.001	0.000	0.000	0.002	0.000	0.000	0.000	0.000	0.000	0.000	0.001	0.000	0.000	0.000	0.000	0.001
e 0.000	0.001	0.001	0.004	0.002	0.001	0.001	0.001	0.000	0.002	0.001	0.006	0.004	0.006	0.000	0.001	0.007	0.003	0.005	0.000	0.000	0.000	0.000	0.001	0.001	0.001	0.000	0.000	0.000	0.000	0.000	0.017
ж 0.001	0.000	0.000	0.000	0.001	0.003	0.000	0.000	0.002	0.000	0.000	0.000	0.000	0.001	0.001	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
з 0.006	0.000	0.001	0.000	0.001	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.002	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.002
и 0.000	0.000	0.002	0.000	0.002	0.001	0.000	0.002	0.000	0.001	0.002	0.006	0.002	0.003	0.000	0.000	0.001	0.003	0.003	0.000	0.000	0.001	0.001	0.001	0.001	0.000	0.000	0.000	0.000	0.000	0.001	0.019
й 0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.007
к 0.008	0.000	0.000	0.000	0.000	0.001	0.000	0.000	0.003	0.000	0.000	0.001	0.000	0.000	0.008	0.000	0.002	0.000	0.001	0.002	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.005
л 0.006	0.000	0.000	0.000	0.000	0.004	0.000	0.000	0.006	0.000	0.000	0.000	0.000	0.000	0.006	0.000	0.000	0.002	0.000	0.001	0.000	0.000	0.000	0.000	0.000	0.000	0.001	0.003	0.000	0.001	0.001	0.011
m 0.003	0.000	0.000	0.000	0.000	0.004	0.000	0.000	0.003	0.000	0.000	0.000	0.000	0.002	0.004	0.000	0.000	0.000	0.000	0.002	0.000	0.000	0.000	0.000	0.000	0.000	0.002	0.000	0.000	0.000	0.001	0.008
н 0.010	0.000	0.000	0.000	0.000	0.009	0.000	0.000	0.006	0.000	0.000	0.000	0.000	0.002	0.010	0.000	0.000	0.000	0.001	0.003	0.000	0.000	0.000	0.000	0.000	0.000	0.003	0.001	0.000	0.000	0.002	0.005
0.000	0.003	0.006	0.005	0.004	0.002	0.002	0.001	0.001	0.004	0.002	0.005	0.005	0.008	0.000	0.001	0.005	0.005	0.007	0.000	0.001	0.001	0.000	0.002	0.001	0.000	0.000	0.000	0.000	0.000	0.001	0.023
n 0.002	0.000	0.000	0.000	0.000	0.003	0.000	0.000	0.001	0.000	0.000	0.001	0.000	0.000	0.011	0.000	0.006	0.000	0.000	0.001	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
p 0.006	0.000	0.000	0.000	0.000	0.005	0.000	0.000	0.005	0.000	0.000	0.000	0.000	0.001	0.006	0.000	0.000	0.001	0.001	0.004	0.000	0.000	0.000	0.000	0.000	0.000	0.001	0.001	0.000	0.000	0.001	0.001
c 0.001	0.000	0.001	0.000	0.000	0.003	0.000	0.000	0.003	0.000	0.003	0.003	0.001	0.001	0.002	0.002	0.000	0.001	0.009	0.001	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.003	0.000	0.000	0.004	0.003
т 0.006	0.000	0.002	0.000	0.000	0.005	0.000	0.000	0.003	0.000	0.001	0.000	0.000	0.001	0.014	0.000	0.003	0.001	0.000	0.002	0.000	0.000	0.000	0.000	0.000	0.000	0.001	0.005	0.000	0.000	0.001	0.006
y 0.001	0.001	0.001	0.001	0.001	0.000	0.001	0.000	0.000	0.000	0.001	0.002	0.001	0.000	0.000	0.001	0.001	0.001	0.002	0.000	0.000	0.000	0.000	0.001	0.001	0.000	0.000	0.000	0.000	0.001	0.000	0.006
ф 0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.001	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.001	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
x 0.001	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.002	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.003
ц 0.000	0.000	0.000	0.000	0.000	0.001	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
4 0.002	0.000	0.000	0.000	0.000	0.003	0.000	0.000	0.001	0.000	0.000	0.000	0.000	0.001	0.000	0.000	0.000	0.000	0.004	0.001	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
w 0.001	0.000	0.000	0.000	0.000	0.002	0.000	0.000	0.001	0.000	0.001	0.001	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.001	0.000	0.000	0.000	0.000
w 0.001	0.000	0.000	0.000	0.000	0.002	0.000	0.000	0.001	0.000	0.001	0.001	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.001	0.000	0.000	0.000	0.000
ы 0.000	0.000	0.001	0.000	0.000	0.001	0.000	0.000	0.000	0.001	0.000	0.002	0.001	0.000	0.000	0.000	0.000	0.000	0.001	0.000	0.000	0.001	0.000	0.000	0.001	0.000	0.000	0.000	0.000	0.000	0.000	0.005
ь 0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.001	0.000	0.000	0.001	0.000	0.000	0.000	0.001	0.000	0.000	0.000	0.000	0.000	0.000	0.001	0.000	0.000	0.000	0.000	0.000	0.000	0.010
э 0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.001	0.000	0.000	0.000	0.000	0.000	0.000	0.003	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
ю 0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.002
я 0.000	0.000	0.000	0.000	0.001	0.000	0.000	0.000	0.000	0.000	0.000	0.001	0.000	0.001	0.000	0.000	0.000	0.000	0.001	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.013
0.002	0.007	0.014	0.003	0.009	0.004	0.001	0.005	0.009	0.000	0.010	0.002	0.008	0.017	0.011	0.018	0.004	0.016	0.008	0.005	0.000	0.002	0.000	0.006	0.001	0.000	0.000	0.000	0.003	0.000	0.005	0.000
H2 with	spaces:	3.980	1036688	866813																											
R for H2	with s	paces:	0.2039	7926622	26637																										

Найчастіші біграми(текс	т з пробілами
та перетином бі	грам)
2.6	0.017

Travisaciimi oli pamii (Teker 3 ripooliiamii											
та перетином бі	грам)										
3 буквою а	_	0,017									
3 буквою б	0	0,002									
3 буквою в	0	0,005									
3 буквою г	Д	0,002									
3 буквою д	е	0,005									
3 буквою е	ı	0,017									
3 буквою ж	e	0,003									
3 буквою з	а	0,006									
3 буквою и	ı	0,019									
3 буквою й	1	0,007									
3 буквою к	0	0,008									
3 буквою л	-	0,011									
3 буквою м	_	0,008									
3 буквою н	Н	0,01									
3 буквою о		0,023									

	_	
3 буквою п	О	0,011
3 буквою р	а	0,006
3 буквою с	Т	0,009
3 буквою т	0	0,014
3 буквою у	_	0,006
3 буквою ф	ф	0,001
3 буквою х	_	0,003
3 буквою ц	е	0,001
3 буквою ч	Т	0,004
3 буквою ш	е	0,002
3 буквою щ	е	0,002
3 буквою ы	_	0,005
3 буквою ь	_	0,01
3 буквою э	Л	0,01
3 буквою ю		0,002
3 буквою я		0,013
3 буквою _	П	0,018

Значення Н2 без пробілів та з перетином:

H2: 4.188308624728646

а	6	В	г	д	е	ж	3	И	й	к	л	м	н	0	п	р	С	т	у	ф	×	ц	ч	ш	ц	ы	ь	3	ю	я
a 0.000	0.002	0.004	0.001	0.003	0.002	0.001	0.005	0.001	0.001	0.008	0.011	0.004	0.005	0.001	0.003	0.003	0.886	0.005	0.001	0.000	0.001	0.888	0.002	0.001	0.000	0.000	0.000	0.000	0.001	0.002
6 0.001	0.000	0.000	0.000	0.000	0.002	0.000	0.000	0.001	0.000	0.000	0.001	0.000	0.000	0.003	0.000	0.002	0.000	0.000	0.001	0.000	0.000	0.000	0.000	0.000	0.000	0.004	0.000	0.000	0.000	0.801
в 0.005	0.000	0.000	0.000	0.001	0.005	0.000	0.001	0.004	0.000	0.001	0.001	0.000	0.002	0.006	0.001	0.001	0.004	0.001	0.001	0.000	0.000	0.000	0.000	0.001	0.000	0.002	0.000	0.000	0.000	0.000
r 0.001	0.000	0.000	0.000	0.002	0.000	0.000	0.000	0.001	0.000	0.000	0.002	0.000	0.000	0.009	0.000	0.001	0.000	0.000	0.001	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
д 0.005	0.000	0.001	0.000	0.000	0.007	0.002	0.000	0.003	0.000	0.000	0.001	0.000	0.002	0.004	0.000	0.001	0.000	0.000	0.002	0.000	0.000	0.000	0.000	0.000	0.000	0.001	0.001	0.000	0.000	0.000
e 0.000	0.002	0.003	0.005	0.004	0.002	0.001	0.002	0.001	0.002	0.003	0.008	0.005	0.009	0.001	0.003	0.009	0.006	0.007	0.001	0.000	0.001	0.000	0.002	0.001	0.001	0.000	0.000	0.000	0.000	0.801
ж 0.002	0.000	0.000	0.000	0.001	0.004	0.000	0.000	0.002	0.000	0.000	0.000	0.000	0.001	0.001	0.000	0.000	0.888	0.000	0.000	0.000	0.000	0.888	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
з 0.007	0.000	0.001	0.000	0.001	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.002	0.001	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.888	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
и 0.000	0.001	0.004	0.001	0.003	0.002	0.000	0.003	0.002	0.001	0.003	0.007	0.004	0.006	0.002	0.003	0.001	0.005	0.005	0.001	0.000	0.002	0.001	0.002	0.001	0.000	0.000	0.000	0.000	0.000	0.802
й 0.000	0.000	0.001	0.000	0.001	0.000	0.000	0.000	0.001	0.000	0.001	0.000	0.001	0.001	0.000	0.001	0.000	0.001	0.001	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
к 0.009	0.000	0.001	0.000	0.000	0.001	0.000	0.000	0.004	0.000	0.000	0.001	0.000	0.001	0.010	0.001	0.003	0.001	0.001	0.002	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
л 0.007	0.001	0.001	0.000	0.001	0.006	0.000	0.001	0.008	0.000	0.001	0.001	0.001	0.002	0.008	0.001	0.000	0.003	0.001	0.002	0.000	0.000	0.000	0.001	0.000	0.000	0.001	0.004	0.000	0.001	0.003
м 0.003	0.000	0.001	0.000	0.001	0.005	0.000	0.000	0.004	0.000	0.001	0.000	0.000	0.003	0.006	0.001	0.000	0.881	0.000	0.003	0.000	0.000	0.000	0.000	0.000	0.000	0.002	0.000	0.000	0.000	0.001
н 0.012	0.000	0.001	0.000	0.001	0.011	0.000	0.000	0.007	0.000	0.001	0.000	0.000	0.003	0.012	0.001	0.000	0.881	0.001	0.004	0.000	0.000	0.888	0.000	0.000	0.000	0.003	0.001	0.000	0.000	0.002
o 0.000	0.005	0.009	0.006	0.006	0.003	0.003	0.002	0.002	0.004	0.004	0.007	0.008	0.012	0.002	0.004	0.007	0.009	0.010	0.001	0.001	0.001	0.000	0.004	0.001	0.000	0.000	0.000	0.001	0.000	0.802
n 0.002	0.000	0.000	0.000	0.000	0.003	0.000	0.000	0.001	0.000	0.000	0.001	0.000	0.000	0.013	0.000	0.007	0.000	0.000	0.001	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
p 0.008	0.000	0.001	0.000	0.001	0.006	0.000	0.000	0.006	0.000	0.001	0.000	0.000	0.001	0.008	0.000	0.001	0.002	0.001	0.005	0.000	0.000	0.000	0.000	0.000	0.000	0.002	0.001	0.000	0.000	0.001
c 0.002	0.000	0.002	0.000	0.001	0.004	0.000	0.000	0.003	0.000	0.004	0.004	0.002	0.002	0.003	0.002	0.000	0.001	0.011	0.001	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.003	0.000	0.000	0.805
T 0.007	0.000	0.003	0.000	0.000	0.006	0.000	0.000	0.004	0.000	0.001	0.000	0.001	0.002	0.018	0.001	0.003	0.002	0.001	0.002	0.000	0.000	0.000	0.000	0.000	0.000	0.002	0.006	0.000	0.000	0.001
y 0.001	0.001	0.001	0.001	0.002	0.000	0.002	0.001	0.001	0.000	0.002	0.003	0.002	0.001	0.001	0.001	0.001	0.882	0.003	0.000	0.000	0.001	0.888	0.001	0.001	0.000	0.000	0.000	0.000	0.001	0.000
ф 0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.001	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.001	0.000	0.888	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
x 0.001	0.000	0.001	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.001	0.003	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
ц 0.001	0.000	0.000	0.000	0.000	0.001	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
ч 0.002	0.000	0.000	0.000	0.000	0.004	0.000	0.000	0.002	0.000	0.001	0.000	0.000	0.001	0.000	0.000	0.000	0.000	0.005	0.001	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
ш 0.001	0.000	0.000	0.000	0.000	0.003	0.000	0.000	0.001	0.000	0.001	0.001	0.000	0.000	0.001	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.001	0.000	0.000	0.000
ш 0.001	0.000	0.000	0.000	0.000	0.003	0.000	0.000	0.001	0.000	0.001	0.001	0.000	0.000	0.001	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.001	0.000	0.000	0.000
ы 0.000	0.000	0.001	0.000	0.000	0.001	0.000	0.000	0.000	0.002	0.000	0.002	0.002	0.001	0.000	0.001	0.000	0.881	0.001	0.000	0.000	0.001	0.888	0.000	0.001	0.000	0.000	0.000	0.000	0.000	0.000
ь 0.000	0.000	0.001	0.000	0.001	0.001	0.000	0.000	0.001	0.000	0.002	0.000	0.001	0.003	0.001	0.001	0.000	0.002	0.001	0.000	0.000	0.000	0.000	0.001	0.001	0.000	0.000	0.000	0.000	0.000	0.801
э 0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.001	0.000	0.000	0.000	0.000	0.000	0.000	0.003	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
ю 0.000	0.000	0.000	0.000	0.001	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.001	0.000	0.001	0.000	0.000	0.000	0.000	0.000
я 0.000	0.001	0.002	0.000	0.002	0.000	0.000	0.001	0.001	0.000	0.001	0.002	0.001	0.003	0.001	0.002	0.000	0.002	0.002	0.001	0.000	0.000	0.000	0.001	0.000	0.000	0.000	0.000	0.000	0.000	0.000
H2 witho	ut spac	es: 4.	1883086	2472864	6																									
R for H2	withou	t space	s: 0.1	6233827	5054270	18																								

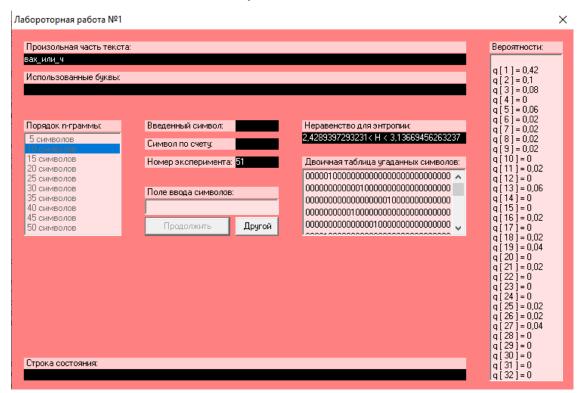
Найчастіші біграми(текст без пробілів та з перетином біграм)												
3 буквою а	С	0,006										
3 буквою б	Ы	0,004										
3 буквою в	0	0,006										
3 буквою г	0	0,009										
3 буквою д	е	0,007										
3 буквою е	Н	0,009										
3 буквою ж	е	0,004										
3 буквою з	а	0,007										
3 буквою и	Л	0,007										
3 буквою й	Д	0,001										
3 буквою к	0	0,01										
3 буквою л	0	0,008										
3 буквою м	0	0,006										
3 буквою н	а	0,012										
3 буквою о	Н	0,012										

3 буквою п	0	0,013
3 буквою р	а	0,008
3 буквою с	Т	0,011
3 буквою т	0	0,018
3 буквою у	Т	0,003
3 буквою ф	И	0,001
3 буквою х	0	0,003
3 буквою ц	а	0,001
3 буквою ч і	Т	0,005
3 буквою ш	е	0,003
3 буквою щ	е	0,003
3 буквою ы	Л	0,002
3 буквою ь	Н	0,003
3 буквою э	Т	0,003
3 буквою ю	Д	0,001
3 буквою я	В	0,002

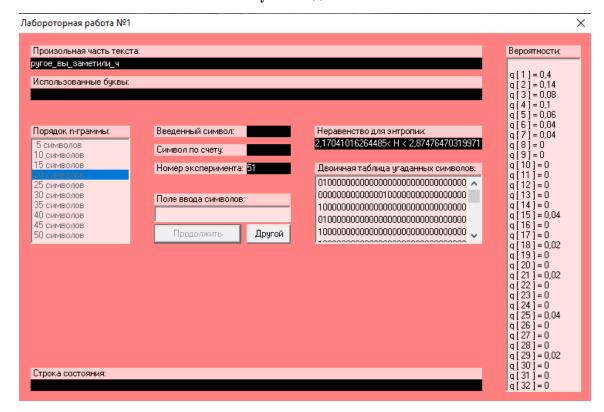
H1 with spaces:	4.366187862436842
R:	0.12676242751263156
H1 without spaces:	4.474445521025726
R:	0.10511089579485478
H2 without crossing and R	
H2 with spaces:	3.978541281663438
R:	0.2042917436673124
H2 without spaces:	4.1880373760474665
R:	0.1623925247905067
H2 with crossing and R	
H2 with spaces:	3.9801036688866813
R:	0.2039792662226637
H2 without spaces:	4.188308624728646
R:	0.1623382750542708

За допомогою програми CoolPinkProgram оцінили значення $H^{(10)}$, $H^{(20)}$, $H^{(30)}$. У кожному досліді було зроблено не менше ніж 50 експериментів.

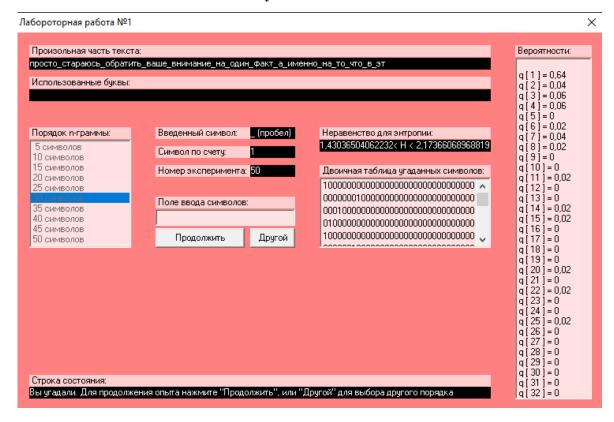
${ m Pe}$ зультат для ${ m H}^{(10)}$



Результат для ${\bf H}^{(20)}$



Результат для H⁽³⁰⁾



Результат надлишковості російської мови

Hadлишковість джерела відкритого тексту (мови) дорівнює $R=1-\frac{H_{\infty}}{H_{0}}$

 $H_0 = log_2 32 = 5$

 $1-(2.428/5)=0.5144>R^{(10)}>1-(3.136/5)=0.3728$

Аналогічні дії повели і для інших випадків.

 $0.566 > R^{(20)} > 0.4252$

 $0.714 > ^{(30)} > 0.5654$

Висновки: під час лабораторної роботи, ми закріпили свої знання стосовно етропії та надлишковості тексту. Вирахували значення для вхідного тексту для монограм та біграм. Удосконалили свої знання мови руthon для роботи з різними структурами даних та файлами, навчилися краще придумовувати алгоритми для роботи з вхідним текстом.